Deep Reinforcement Learning Based Approach for Proportional Resonance Power System Stabilizer to Prevent Ultra-Low-Frequency Oscillations

Zhang, Guozhou; Hu, Weihao; Cao, Di; Huang, Qi; Yi, Jianbo; Chen, Zhe; Blaabjerg, Frede

Published in:
I E E E Transactions on Smart Grid

DOI (link to publication from Publisher):
10.1109/TSG.2020.2997790

Publication date:
2020

Document Version
Accepted author manuscript, peer reviewed version

Deep Reinforcement Learning Based Approach for Proportional Resonance Power System Stabilizer to Prevent Ultra-Low-Frequency Oscillations

Guozhou Zhang, Weihao Hu, *Senior Member, IEEE,* Di Cao, Qi Huang, *Senior Member, IEEE,* Jianbo Yi, *Member, IEEE,* Zhe Chen, *Fellow, IEEE,* Frede Blaabjerg, *Fellow, IEEE*

***Abstract*— Recent studies have shown that due to the hammer effect of the governor, hydropower units are easily creating negative damping torque at the common mode frequency (below 0.1 Hz). Therefore, there is a risk of ultra low frequency oscillations (ULFO) in hydropower-dominated systems. ULFO is a small-signal frequency oscillation problem, which is quite different from low frequency oscillations (LFO). A conventional power system stabilizer (CPSS) has less effect on suppressing ULFO. To solve this problem, this paper proposes a high-order polynomial structure to replace the CPSS, and combine it with a proportional resonance controller to form a novel PR-PSS. In order to ensure the robustness of PR-PSS, based on the characteristic analysis results of the PR-PSS, a deep reinforcement learning (DRL) algorithm asynchronous advantage actor-critic (A3C) is introduced to train an agent. After training, the proposed agent can provide optimal parameter settings for PR-PSS under various operating conditions. Simulation results verify the effectiveness of the proposed method.**

***Index Terms*— ultra-low-frequency oscillations, low frequency oscillation, CPSS, PR-PSS, deep reinforcement learning.**

I. Introduction

In recent years, frequency oscillations below 0.1 Hz have occurred several times, especially in power systems with a high proportion of hydropower. In 2016, when engineers tested asynchronous operation in Yunnan Power Grid (China Provincial Power Grid), the system had an oscillation with a frequency of 0.05 Hz and the frequency fluctuated between 49.9 Hz and 50.1 Hz [1]. Similar phenomena were observed in Turkey and Colombia [2-3]. The simulation results of Southwest China Power Grid (SCPG) showed that over 80% of N-1 faults and 60% of N-2 faults in the system would activate ultra-low-frequency oscillations (ULFOs) [4].

After careful studies, several papers concluded that the above ULFO events are closely related to primary frequency regulation (PFR) of the hydraulic turbine. More specifically, the setting of hydraulic governor for quick dynamic response easily leads to ULFO owing to the water hammer effect [5-6]. Ref [7] adopted a damping torque analysis method to analyze the damping of the hydraulic turbine-governor and found out that it easily creates negative damping torque.

Various approaches have been proposed in the literature to suppress ULFO. Studies in [5] show that, unlike hydraulic turbines, steam turbines can provide positive damping in ultra-low frequency band (ULFB). Ref. [8] suggested that the ULFO can be suppressed by increasing the proportion of thermal power. However, this approach would reduce the capability of hydro energy accommodation, and affect the peak shaving and frequency control of the power system. In [6], the authors proposed a method to suppress ULFO by optimizing the parameters of high voltage direct current frequency limiting controller (HVDC-FLC). However, in some hydro-dominant systems, electric power is transmitted only through alternating current (AC) transmission. Therefore, it is not feasible for these regions to suppress ULFO by adjusting FLC parameters.

Re-tuning the governor settings is also proposed to suppress ULFO. In [9], the PI parameters of the governor were optimized to improve the damping torque. Authors in [10] suggested quitting the PFR with negative damping as an emergency strategy. In this way, ULFO can be suppressed. Although these strategies can suppress ULFO, they will reduce the dynamic response of the governor.

A conventional power system stabilizer (CPSS) is an effective device to damp the low frequency oscillation (LFO). Authors in [7] attempted to suppress ULFO with CPSS. However, simulation results show that CPSS cannot eliminate ULFO. The reason is that CPSS provides a larger phase lead in ULFB, which limits the positive damping provided by CPSS [11]. In fact, in addition to the excitation system (CPSS), PSS can also be attached to the governor (GPSS). In [12], the authors applied GPSS to suppress ULFO and simulation results show that the damping of ULFO mode can be improved by GPSS. However, authors in [12] ignored the impact of GPSS on the LFO modes. The damping of the LFO modes may decrease in the process of suppressing ULFO. Moreover, the experimental results showed that the governor oil pumps will have serious overheating problems due to the long-term operation of GPSS [13]. A multi-band PSS is proposed by Hydro-Québec, which is named PSS4B [14]. The PSS4B consists of three bands, one of the bands can

This work was supported by The National Science Fund for Distinguished Young Scholars of Sichuan Province number 20JCQN0213 and National Key Research and Development Project of Sichuan Province number 20ZDYF0458 (Corresponding author: Weihao Hu)

Guozhou Zhang, Weihao Hu, Di Cao, Qi Huang, Jianbo Yi are with the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: zgz@std.uestc.edu.cn; whu@uestc.edu.cn; caodi@std.uestc.edu.cn; hwong@uestc.edu.cn; jimbo_yi@uestc.edu.cn)

Zhe Chen and Frede Blaabjerg are with the Department of Energy Technology, Aalborg University, Aalborg, Denmark (e-mail: zch@et.aau.dk, fbl@et.aau.dk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

theoretically provide damping for ULFO modes. However, PSS4B is rather rigid when it is applied to suppress ULFO and further investigations are needed [7].

To solve the above problems, this paper combines a high-order polynomial structure with proportional resonance to form a novel PR-PSS, which overcomes the shortcomings of CPSS in suppressing ULFO. The characteristic analysis of PR-PSS shows that it has less influence on other eigenvalues, and it does not reduce the damping of the LFO modes in the process of suppressing ULFO.

For the PR-PSS, the parameter settings play an important role in suppressing the ULFO. To obtain the optimized parameter settings, various methods have been applied to optimize the controller in previous studies. In [15], the residue analysis method is applied to tune the damping controller. In [16], the particle swarm optimization (PSO) algorithm is used for PSS parameter optimization. Moreover, the robust design method, e. g. fuzzy theory [17], H- ∞ control [18] and eigenvalue sensitivity [19] are also used to design a robust PSS.

In recent years, with the development of artificial intelligence (AI), deep reinforcement learning (DRL) has become a promising alternative with better robustness to tune the controller. In [20], a Q-learning algorithm is used for adaptive adjustment of PSS. However, Q-learning is only applicable for scenarios that have a low-dimensional and discrete action domain. When it is used to design PSS, the action domain needs to be discretized, which weakens the control effect. To solve this problem, an asynchronous advantage actor-critic (A3C) algorithm is proposed in [21] and it is suitable for scenarios with continuous action domain. In [22], the A3C algorithm is applied to solve the energy management problem.

In this paper, to ensure the effectiveness of the PR-PSS in various operation conditions, an A3C algorithm is introduced into adaptive control of PR-PSS. The contributions of this paper are summarized as follows:

1) Based on a damping torque coefficient method, the mechanism of ULFO is studied, and the differences between LFO and ULFO are discussed. Moreover, the feasibility of suppressing ULFO by PSS is analyzed.

2) To overcome the shortcomings of traditional PSS, a novel PR-PSS is proposed to suppress ULFO, which is a novel approach to solve this problem. In contrast with other classical PSS tuned devices, it can provide better damping in the ultra-low frequency band and suppress ULFO more effectively.

3) The adaptive adjustment of the PR-PSS is formed as a Markov decision process (MDP) with finite time steps, and the A3C algorithm is introduced to solve the MDP to obtain the optimal policy. Which ensures the effectiveness of PR-PSS in each operating condition.

4) Cases studies demonstrate that the proposed PR-PSS controller can show better performance compared with other ULFO suppression strategies. Moreover, the proposed agent can provide optimal parameters setting for the PR-PSS in each operating condition, and make the system has a larger stability margin than traditional parameter setting method.

The remaining part of this paper is organized as follows: The mechanism of ULFO is presented in section II. In section III, the structure of PR-PSS is introduced. Section IV analyses the character of PR-PSS and proposes the parameters self-tuning method for PR-PSS. Simulation results are provided in Section V and Section VI. Section VII concludes the paper.

II. MECHANISM OF ULFO

In this section, the Phillips-Heffron [23] model including turbine governing system and PSS are constructed to analyze the mechanism of ULFO.

The standard linear model of Phillips-Heffron is shown in Fig. 1, and the motion equation of generator rotor considering only the small signal is as follows:

$$\begin{cases} \Delta\delta = \dfrac{\omega_0}{s}\left[\dfrac{1}{Ms+D}\left(\Delta T_m - \Delta T_e\right)\right] \\ \Delta\omega = \dfrac{s\Delta\delta}{\omega_0} \end{cases} \tag{1}$$

where $\Delta\delta$ and $\Delta\omega$ are the increment of rotor angle and rotor speed, respectively; M is inertia time constant of the generator; ΔT_m and ΔT_e are the increment of mechanical torque and electromagnetic torque, respectively. D is the damping coefficient. In general, it can be neglected [7].

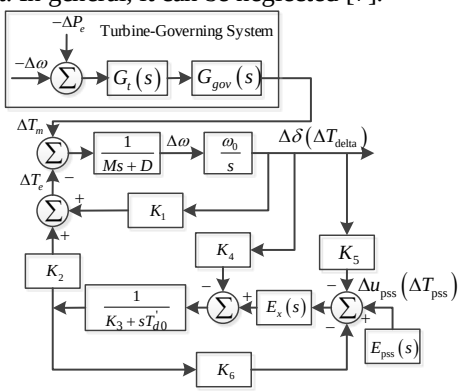


Fig. 1. Linearized model of Phillips-Heffron [23] in a connected network.

Based on the damping torque analysis method, the mechanical torque increment ΔT_m can be decomposed into:

$$\Delta T_m = -K_{md}\Delta\omega - K_{ms}\Delta\delta \tag{2}$$

where K_{md} and K_{ms} are the damping torque coefficient and the synchronous torque coefficient of ΔT_m, which can be calculated as follows:

$$\begin{cases} G_m(s) = G_{\text{gov}}(s)G_t(s) \\ K_{md} = \text{Re}(G_m(j\omega_d)) \\ K_{ms} = \text{Im}(G_m(j\omega_d)) \end{cases} \tag{3}$$

$G_{\text{gov}}(s)$ represents the governor model. In this paper, the simplified governor model is discussed and the transfer function of governor is [5]:

$$G_{gov}(s) = \frac{1}{T_G s + 1} \tag{4}$$

where T_G is the response time of the governor.

$G_t(s)$ represents the turbine model and the transfer function is listed as follows:

$$G_t(s) = \frac{1 - T_w s}{1 + 0.5T_w s} \tag{5}$$

where T_w is time constant of water hammer effect.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

Substituting $s = jw_d$ into the equation (4) and equation (5), it can be obtained that:

$$K_{md} = \frac{1-0.5T_G T_w \omega_d^2 - \omega_d^2 T_w \left(T_G + 0.5T_w\right)}{\left(1-0.5T_G T_w \omega_d^2\right)^2 + \omega_d^2 \left(T_G + 0.5T_w\right)^2} \tag{6}$$

For equation (6), set T_G = 5, and bring $w_d = 2\pi f$ into equation (3) to obtain the trajectory of K_{md} change with frequency f under different T_w conditions. The results are shown in Fig. 2.

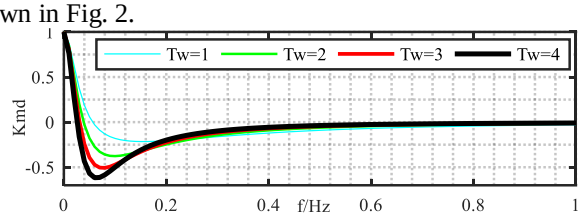


Fig. 2. Damping characteristic curve under a variety of T_w conditions.

It can be seen from Fig. 2 that the damping coefficient K_{md} is negative in the ULFB (<0.1 Hz). Moreover, with an increase of T_w, the K_{md} would decrease.

The electromagnetic torque increment ΔT_e of the generator can be decomposed into:

$$\Delta T_e = \Delta T_{\text{delta}} + \Delta T_{\text{pss}} \tag{7}$$

where ΔT_{pss} is the torque increment contributed by the output signal of the PSS; ΔT_{delta} is the torque increment contributed by excitation system.

Based on the damping torque analysis method, ΔT_{delta} and ΔT_{pss} are decomposed into damping torque (Δw) and synchronization torque ($\Delta\delta$) respectively:

$$\begin{cases} \Delta T_{\text{delta}} = K_{\text{delta-}d}\Delta\omega + K_{\text{delta-}s}\Delta\delta \\ \Delta T_{\text{pss}} = K_{\text{pss-}d}\Delta\omega + K_{\text{pss-}s}\Delta\delta \end{cases} \tag{8}$$

where $K_{\text{delta-d}}$ and $K_{\text{pss-d}}$ are damping torque coefficients of ΔT_e, respectively; $K_{\text{delta-s}}$ and $K_{\text{pss-s}}$ are synchronous torque coefficients of ΔT_e. Then, the following equation can be derived from equation (7) and (8):

$$\begin{aligned} &\Delta T_e = K_{ed}\Delta\omega + K_{es}\Delta\delta \\ &\begin{cases} K_{ed} = K_{\text{delta-}d} + K_{\text{pss-}d} \\ K_{es} = K_{\text{delta-}s} + K_{\text{pss-}s} \end{cases} \end{aligned} \tag{9}$$

Based on (2) and (9), neglecting D, the motion equation of generator (1) can be further formulated as:

$$s^2(\Delta\delta) + \frac{1}{M}\left(K_{ed} + K_{md}\right)s(\Delta\delta) + \frac{\omega_0}{M}\left(K_{es} + K_{ms}\right)(\Delta\delta) = 0 \tag{10}$$

The damping ratio ξ is as follows:

$$\xi = \frac{1}{2\sqrt{M\omega_0}} \frac{K_{ed} + K_{md}}{\sqrt{K_{es} + K_{ms}}} = \frac{1}{2\sqrt{M\omega_0}} \frac{K_{\text{delta-}d} + K_{\text{pss-}d} + K_{md}}{\sqrt{K_{es} + K_{ms}}} \tag{11}$$

In fact, for the low frequency oscillation (LFO) or electro-mechanical oscillation, it is caused by the negative damping of excitation system. More specially, due to the wide application of the high magnification excitation system, the excitation system easily make the damping torque coefficients $K_{delta\text{-}d}$ negative. Moreover, this negative $K_{delta\text{-}d}$ would make the system appear to have a negative damping oscillation mode in the range of 0.1-2 Hz (named LFO mode).

For the ULFO problem, it is caused by the negative damping of the hydraulic governor. More specifically, due to the water hammer effect, the setting of hydraulic governor for quick dynamic response can easily produce negative damping torque K_{md} in the ULFB (Fig. 2), which makes the system appear to have a negative damping oscillation mode in the range of 0.01-0.1 Hz (named ULFO mode).

ULFO is completely different from the LFO. No oscillations between units are observed. In the frequency oscillations, the speed of generators and the frequencies of buses vary with the same phase and amplitude [6]. Moreover, it is strongly related with governors and turbines of generators, and is deemed as a result of the small-signal instability of the system's primary frequency control process. The ULFO is not a problem of angle stability, but belongs to the category of frequency stability.

The reason why CPSS can prevent LFO is that the damping torque coefficients $K_{\text{delta-d}}$ (which is caused by the excitation system) can be offset by the $K_{\text{pss-d}}$ (which is caused by the CPSS). This makes the damping of the system positive.

Theoretically, the negative damping torque K_{md} can also be offset by a positive damping torque $K_{\text{pss-}d}$ produced by CPSS and the ULFO can be avoided in this way.

However, CPSS cannot provide enough damping torque in ULFB. Authors in [11] do various experiments to analyze the amplitude and phase characteristics of CPSS. The results show that CPSS always provides an excessive phase lead in ULFB, and this makes CPSS hardly to provide damping torques for both LFO modes and ULFO modes. Similarly, the simulation results in [7] also show that CPSS cannot suppress ULFO. Hence, it is not feasible to suppress ULFO by configuring CPSS.

III. STRUCTURE OF PR-PSS

As mentioned in Section II, the reason why CPSS cannot suppress ULFO is that it can only provide a slight damping in ULFB. If this shortcoming can be overcome by improving its structure, PSS can be used to prevent ULFO. Based on the above considerations, a PR-PSS is introduced in this section.

A. PSS Model

For a PSS, previous studies have shown that the structure of PSS will affect the control effect. In this paper, to ensure that the optimal PSS structure can be obtained to suppress ULFO, a high-order polynomial is proposed to replace the traditional PSS structure and it is listed as:

$$G_{PSS}(s) = \frac{a_0 s^n + a_1 s^{n-1} + \cdots + a_n}{b_0 s^n + b_1 s^{n-1} + \cdots + b_n} \tag{12}$$

In fact, equation (12) can be regarded as a uniform expression of the PSS transfer function. By adjusting the parameters of equation (12), an optimal structure of PSS can be obtained to suppress ULFO.

B. PR Controller Model

The proportional resonance controller consists of a proportional model and a resonant model. The main characteristics of a PR controller is that it can accurately control one certain frequency and reject the other frequencies effectively. The transfer function of a PR controller is as follows [24]:

$$G_{\text{PR}}(s) = K_p + \frac{2K_R\omega_c s}{s^2 + 2\omega_c s + \omega_0^2} \tag{13}$$

where K_P is the proportionality coefficient; K_R is the resonance coefficient; ω_0 is the resonant frequency. ω_c is the cut-off frequency of the resonant model.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

C. PR-PSS Model

To overcome the shortcoming of CPSS, this paper adds the PR controller to the PSS to form a novel PSS, named PR-PSS. The structure of the PR-PSS is shown in Fig. 3. In PR-PSS, the PR controller is set as a band-pass filter, which is used to retain the concerned mode (ULFO mode) by filtering out other modes in the original signal and to obtain a pure sole modal signal. Then, this processed signal is used as the input to the PSS. Compared with the PSS4B, PR-PSS has these advantages:

Fig. 3. Structure of proportional resonance PSS.

1) PR controller can accurately control one certain frequency and reject the other frequencies effectively, and the command signal at this frequency can be tracked without static error. Therefore, this paper introduces a PR mechanism to focus the PSS controller's effect on the target frequencies, and then add damping mainly to that particular frequency.

2) The structure of PR-PSS is much simpler than PSS4B and it has fewer controller parameters. Therefore, it is easier to design the PR-PSS parameters than PSS4B. Moreover, a high-order polynomial is proposed to replace the traditional PSS structure. By adjusting the parameters of high-order polynomial, the optimal structure can be obtained to suppress ULFO.

3) For the CPSS, PSS4B and PR-PSS, the corresponding amplitude-frequency characteristic curves and phase-frequency characteristic curves are shown in Fig. 4. It can be seen that, by adjusting the parameters of the PR-PSS controller, the phase lead amplitude is reduced and the gain is increased. Moreover, PR-PSS can provide better damping in the ULFB compared to both PSS4B and CPSS.

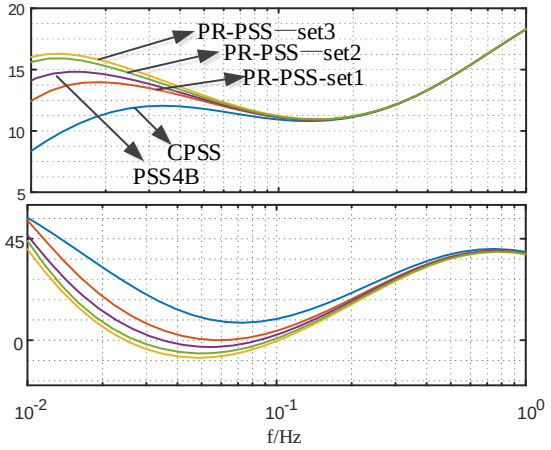


Fig. 4. Bode diagram of PSS, PSS4B and PR-PSS at very low frequency.

IV. PR-PSS CHARACTERISTICS ANALYSIS AND PARAMETER TUNING RULE

The structure of the PR-PSS mentioned in Section III is different from the CPSS. The traditional parameter tuning method may not be so good. Therefore, it is necessary to analyze the characteristics of the PR-PSS. Based on the analysis results, the parameter setting rules are formulated.

A. PR-PSS Characteristics Analysis

The linear state space model of *n*-machine system is as follows [15]:

$$\begin{cases} \dot{x} = Ax + Bu = Ax + \sum_{j=1}^{n} B_j u_j \\ y_j = C_j x \end{cases} \tag{14}$$

where vector x is the state variable of the system; A is the state matrix; vector u_j and y_j are the input and output of the jth generator of the system; B_j and C_j are the input and output matrices, respectively.

For the state matrix A, the following equation can be obtained:

$$\begin{aligned} AM &= M\Lambda \\ MN^T &= I \end{aligned} \tag{15}$$

where M is right feature vector; N is left feature vector. Λ is the diagonal matrix, which can be expressed as follows:

$$\Lambda = \mathrm{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_m\right) \tag{16}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of the state matrix A.

Given the following transformation:

$$x = Mz \tag{17}$$

By introducing the above equation into the linear state space model of the system, it can be obtained that:

$$\begin{cases} \dot{z} = M^{-1}AMz + M^{-1}\sum_{j=1}^{n} B_j u_j = \Lambda z + N^T \sum_{j=1}^{n} B_j u_j \\ y_j = C_j Mz = \sum_{i=1}^{n} C_j m_i z_i \end{cases} \tag{18}$$

Since y_j passes through PR band-pass filter, it is only a single mode signal: $y_j = C_j m_i z_i$. Where z_i is the ith component of z. The output of PR-PSS configured on the jth generator is as follows:

$$u_j = G_{PR\text{-}PSS\,j}(s)\, y_j = G_{PR\text{-}PSS\,j}(s) C_j m_i z_i \tag{19}$$

By introducing equation (19) into the linear state space, it can be obtained that:

$$\dot{z} = \Lambda z + N^T \sum_{j=1}^{n} B_j G_{PR\text{-}PSS\,j}(s) C_j m_i z_i = A^* z \tag{20}$$

A^* is the state matrix of the transformed system, which is constructed as follows:

$$A^* = \begin{bmatrix} \lambda_1 & \cdots & n_1^T \sum_{j=1}^{n} B_j G_{PR\text{-}PSS\,j}(s) C_j m_i & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_i + n_2^T \sum_{j=1}^{n} B_j G_{PR\text{-}PSS\,j}(s) C_j m_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & n_n^T \sum_{j=1}^{n} B_j G_{PR\text{-}PSS\,j}(s) C_j m_i & \cdots & \lambda_m \end{bmatrix} \tag{21}$$

The eigenvalues of A^* are:

$$\Lambda^* = \mathrm{diag}\left(\lambda_1, \cdots \lambda_i + n_1^T \sum_{j=1}^{n} B_j G_{PR\text{-}PSS\,j}(s) C_j m_i, \cdots, \lambda_m \right) \tag{22}$$

It can be seen from the above equation that PR-PSS is only valid for a single mode (ULFO mode) obtained by filtering and it has no impact on the other oscillation modes. In fact, PSS has the ability to enhance the damping of LFO modes, and the essence of PR-PSS is to concentrate this capability on

enhancing the damping of ULFO mode.

B. PR-PSS Parameters Self-Tuning Method

Assuming the *i*th eigenvalue λ_i is ULFO mode:

$$\begin{aligned}\lambda_i &= \alpha_i + j\omega_i \\ \xi_i &= -\alpha_i \Big/ \sqrt{(\alpha_i)^2 + (\omega_i)^2}\end{aligned} \tag{23}$$

where α_i and ω_i are respectively the real part and imaginary part of ULFO mode. ξ_i is the damping ratio of ULFO mode.

Studies in [7] show that λ_i is strongly related to the operating conditions of hydraulic turbines and would change with the operating conditions of hydraulic turbines. This means that under different operating conditions, λ_i are different. In order to ensure the validity of the PR-PSS in various operating conditions, this paper introduces an A3C algorithm [21] to learn a policy for an agent to provide the optimal parameter settings during different operating conditions.

1) Problem Formulation

The adjustment of PR-PSS settings is a decision making problem in unknown environments, which can be formulated as a MDP with finite time steps. In general, the MDP can be described by *S*, *A*, *P*, *R*.

• *S* is the state set, as mentioned above, ULFO mode is strongly related to the operating conditions of hydraulic turbines. To make the uncertainty of the system to be better represented, some measured electrical parameters (such as node voltage) in each operating condtions are taken as states. Therefore, the state of the system at the *k*th step can be defined as $s_k = \left(M_{1,k} \cdots M_{i,k} \cdots M_{m,k}\right)$. $M_{i,k}$ is the *i*th bus measured electrical parameter during the *k*th time step.

• A is the action set, in this paper, where action refers to the adjustment of PR-PSS parameters. Therefore, the action set should be PR-PSS parameters set. Noted that the PR controller should be set as a band pass filter to ensure accurate filtering of UFLO mode. The resonant frequency of the PR controller can be set as the center frequency: $\omega_0 = \omega_i$ (ω_i can be identified by Prony algorithm under various operating conditions.).

Considering that there may be an error between the identified ULFO model and its real value, a certain bandwidth is essential. For PR controller, the filter bandwidth of PR controller is determined by K_P and ω_c, and K_R has no effect on it. Set K_P=0 and s=$j\omega$, and bring them into equation (13):

$$G(j\omega) = \frac{2K_R\omega_c j\omega}{-\omega^2 + 2\omega_c j\omega + \omega_0^2} = \frac{K_R}{1 + j\dfrac{\omega_0}{2\omega_c}\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \tag{24}$$

According to the definition of bandwidth, two solutions of equation $\left|G(j\omega)\right| = K_R/\sqrt{2}$ are the cut-off frequencies. It means that the equivalent solutions of $\left|\dfrac{\omega_0}{2\omega_c}\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)\right| = 1$ are the cut-off frequency. The bandwidth is $B = \omega_c/\pi$ and the ω_c of PR controller can be set as $B\pi$. For the setting of K_p and K_R, this paper sets $G_{\mathrm{PR}}(j\omega_0) = 1$. This means that the single mode signal obtained by filtering, amplification or attenuation should be avoided. The mathematically derivation between K_p and K_R can be obtained in this way by: K_P+K_R=1.

It can be concluded that for the settings of the PR controller, only two parameters *B* and K_p need to be adjusted, other parameters can be identified by Prony algorithm (ω_0) or calculated by these two parameters (ω_c, K_R). Therefore, the action a_t at the *k*th step can be defined as $a_k = \left(K_{p,k}, B_k, a_{0,k} \cdots a_{n,k}, b_{0,k} \cdots b_{n,k}\right)$. Where $K_{p,k}$ and B_k are the PR controller parameters setting at the *k*th time step; $a_{0,k} \cdots a_{n,k}, b_{0,k} \cdots b_{n,k}$ are the PSS controller parameters settings at the *k*th time step. In this paper, *n* is set to 2.

• *P* is the state transition probability function and it describes the probability of the environment moving from the state s_k to the next state s_{k+1}. It can be defined as $s_{k+1} \sim P(s_k, a_k)$.

• *R* is the reward function and it can be used to evaluate the merits of the action. r (s_k, a_k) can be used to describe the reward obtained by the agent when taking action a_k at state s_k. In this paper, the reward is used to evaluate the effect of PR-PSS parameter settings on the damping of oscillation modes. As mentioned in Section IV-A, PR-PSS has less effect on other oscillation modes. When adjusting the PR-PSS parameters, only the change of ULFO mode need to be considered. Therefore, the reward at time step *k* can be defined as follows:

$$r(s_k, a_k) = \begin{cases} 0 & \text{if } \xi_{i,k} \geq \xi_{set} \\ -\left|\xi_{set} - \xi_{i,k}\right| & \text{if } \xi_{i,k} < \xi_{set} \end{cases} \tag{25}$$

where ξ_{set} is the desired damping ratio of ULFO mode, it can be set to be 5% [19].

At each time step *k*, the agent perceives the state of the environment s_k and takes an action a_k based on a policy $\pi(a_k|s_k)$ (π is the policy which maps states to actions). Then the environment transfers to the next state s_{t+1} with probability $P(s_k, a_k)$, and this agent obtains a immediate reward r_k. The key to solve the problem of adaptive adjustment of PR-PSS parameters is to obtain the optimal policy π which can make the agent to obtain the maximize discounted reward from the start state. In general, the action-value function is used to map state-action pairs (s_k, a_k) to the expected cumulative discounted reward following the policy π, which can be defined as:

$$Q^{\pi}(s_k, a_k) = \mathbb{E}_{\pi}\left[R_k | s_k, a_k\right] \tag{26}$$

where

$$R_k = r_k(s_k, a_k) + \gamma r_{k+1}(s_{k+1}, a_{k+1}) + \cdots = \sum_{i=k}^{\infty} \gamma^{i-k} r_i(s_i, a_i) \tag{27}$$

where R_k is the cumulative discounted reward from the time step *k*; γ is a discount factor.

Since equation (26) satisfies the recursive relationship, it can be transformed into the Bellman expectation equation [25]:

$$Q^{\pi}(s_k, a_k) = \mathbb{E}_{\pi}\left[R_k + \gamma \mathbb{E}_{a_{k+1} \sim \pi}\left[Q^{\pi}(s_{k+1}, a_{k+1})\right]\right] \tag{28}$$

In order to obtain the optimal policy π for the agent, a start of the art DRL algorithms A3C is introduced in this paper.

2) A3C Algorithm for PR-PSS Adaptive Control

The A3C algorithm is an actor-critic structured based method of the DRL. Both the actor network and the critic network are introduced to approximate the policy function and action-value function, respectively. During the training, these two networks are trained against each other to make the critic network evaluate policy function π more accurate, and the actor network provide better parameter settings for PR-PSS.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

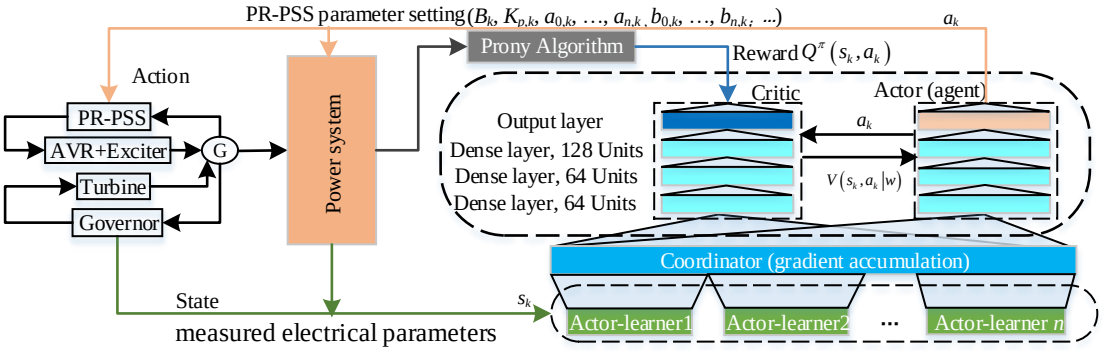


Fig. 5. A3C-based agent training diagram to be applied.

Moreover, asynchronous multi-threaded mechanism is adopted in the A3C algorithm. Multiple threads-specific parallel actor-learners execute different exploration policies, and the overall changes are accumulated to update the networks. The training process is as follows:

For the critic network, it is parameterized by w, takes as input state s_t and action a_t, outputs an estimate of the action-value function $V(s_k, a_k|w)$. The training of critic networks is the process of minimizing the loss function between the output of critic network and the action-value function:

Algorithm 1 Policy learning based on A3C Algorithm

Input: Measured electrical parameters

Output: PR-PSS parameter settings

1: **Initialize:**
2: Assume global shared parameter vectors w', u' and global shared counter $T = 0$
3: Assume thread-specific parameter vectors w, u
4: Initialize thread step counter $t \leftarrow 1$
5: **repeat**
6: Reset gradients: $\nabla_{w'} L(\theta^{w'}) \leftarrow 0$ and $\nabla_{u'} J(\theta^{u'}) \leftarrow 0$
7: Synchronize thread-specific parameters $w'=w$ and $u'=u$.
8: $t_{start} = t$
9: Get state s_t
10: **repeat**
11: Perform a_t based on the policy $\pi(a_t|s_t, \theta^u)$
12: Receive reward r_t and new state s_{t+1}
13: $t \leftarrow t+1$, $T \leftarrow T+1$
14: **until** terminal s_t **or** $t - t_{start} = = t_{max}$
15: $R = \begin{cases} 0 & \text{for terminal } s_t \\ V(s_t, a_t|\theta^w) & \text{for non-terminal } s_t \end{cases}$
16: **for** $i \in \{t-1, \ldots, t_{start}\}$ **do**
17: $R \leftarrow r_i + \gamma R$
18: Accumulate gradients with respect to w':
$$\nabla_{w'} L(w') \leftarrow \nabla_{w'} L(w') + \nabla_{w'} L(w)$$
19: Accumulate gradients with respect to u':
$$\nabla_{u'} J(u') \leftarrow \nabla_{u'} J(u') + \nabla_{u'} J(u)$$
20: **end for**
21: Perform asynchronous update based on (30) and (34)
22: **Until** $T > T_{max}$

$$L(w) = \mathbb{E}_w\left[Q^{\pi}(s_k, a_k) - V(s_k, a_k|w)\right]^2 = \mathbb{E}_w\left[r_k + \gamma V(s_{k+1}, a_{k+1}|w) - V(s_k, a_k|w)\right]^2 \quad (29)$$

In the A3C algorithm, n-steps return method is adopted to improve the training speed of neural networks and the loss function can be rewritten as [21]:

$$L(w) = \mathbb{E}_w\left[\sum_{i=0}^{n-1} \gamma^i r_{k+i} + \gamma^n V(s_{k+1}, a_{k+1}|w) - V(s_k, a_k|w)\right]^2 \quad (30)$$

$$w_{k+1} = w_k + n_w \nabla_w L(w)$$

where n_w is the learning rate of the critic networks.

For the actor network, it is parameterized by u, takes as input state s_t and outputs the action a_t. The actor networks performs the policy $\pi(a_k|s_k, u)$ improvement task, and it updates the policy based on the action-value function estimated by critic networks. To obtain the optimal policy, during the training, the actor networks can be optimized by the policy gradient [22]:

$$\nabla_u J(u) = \mathbb{E}_u\left(\nabla_u \log \pi(a_k|s_k, u)(R_k - V(s_k, a_k|w))\right) \quad (31)$$

Similar to the critic network, n-steps return method is also adopted to the training of the actor network and the policy gradient can be rewritten as:

$$\nabla_u J(u) = \mathbb{E}_u\left(\sum_{t=1}^{T} \nabla_u \log \pi(a_k|s_k, u) A(s_k, a_k)\right) \quad (32)$$

where

$$A(s_k, a_k) = \sum_{i=0}^{n-1} \gamma^i r_{k+i} + \gamma^n V(s_{k+n}, a_{k+n}|u) - V(s_k, a_k|w) \quad (33)$$

In A3C, to explore the environment more effectively, the entropy of the policy function $\nabla_{u'} H(\pi(a_k|s_k, u))$ is added to the policy function and the policy function can be rewritten as:

$$\nabla_u J(u) = \mathbb{E}_u\left(\sum_{t=1}^{T} \nabla_u \log \pi(a_k|s_k, u) A(s_k, a_k) + \beta \nabla_u H(\pi(a_k|s_k, u))\right)$$

$$u_{k+1} = u_k + n_u \nabla_u J(\theta^u) \quad (34)$$

where n_u is the learning rate of the actor networks; H is the entropy; β is the core parameter, which is used to control the strength of the entropy regularization term.

The procedure of the A3C algorithm is outlined in Algorithm1 , and the computational process of the A3C-based agent is shown in Fig. 5.

V. SIMULATION AND DISCUSIONS

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

A. Layout of Testing Case

In this paper, simulations are carried out on a hydropower system in SCPG. The simplified system topology is shown in Fig. 6. The system is representative of the 2016 summer grid, which with 53 busses, 26 lines, and 17 hydropower units. The total installed capacity of the system is 1100 MW. All the units adopt a fifth-order model and they are equipped with the excitation system and turbine governing system. The Prony algorithm is applied to system identification, and the results show that the test system has ULFO mode. Hence, PR-PSS is configured on the CTHC-2#, CPQ-3#, CJB-1# and CJB-2#, which are strongly related with this ULFO mode.

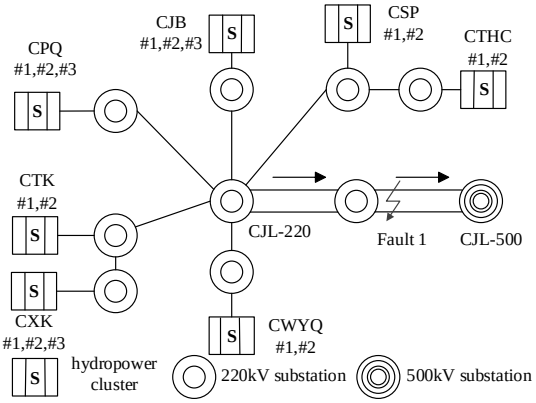


Fig. 6. Transmission network used for case study.

For the test system, the eigenvalues of the test system is calculated, and the results show that there are 16 electro-mechanical oscillation modes in the system. For these modes, the damping of four modes is small and can be regarded as the dominant modes (three LFO modes and one ULFO mode), which are listed in Table I.

Table I
Dominant modes of the test system

Modes	Eigenvalues	Damping (%)	Frequency (Hz)
ULFO mode	-0.012+j0.57	2.10	0.09
LFO mode 1	-0.686+j11.77	5.81	1.87
LFO mode 2	-0.615+j10.28	5.97	1.64
LFO mode 3	-0.454+j9.49	4.78	1.51

B. Training Process of A3C-Based Agent

Before the training, the structure of the actor networks and critic networks is predetermined. In this study, they both adopt three hidden layers, consisting of 128 neurons, 64 neurons and 64 neurons (see Fig. 5), respectively. Besides, the hyper parameters in the A3C algorithm are shown in Table II:

TABLE II
Hyper Parameters in A3C Algorithm

Parameter	Value
Maximum time step (t_{max})	10
Reward discount factor (γ)	0.9
Entropy regularization term (β)	0.01
Learning rate for actor network (η_w)	0.0001
Learning rate for critic network (η_u)	0.00001
Maximum training episode ($T_{\max}$)	10000

During the training process, the history state data is introduced to train the agent. In each episode, the agent would interact with the system to randomly sample a 10 consecutive state set from history data. Then the agent would provide the action based on the state. After that, the states and actions are transferred to the test system, and the time domain simulation and Prony identification are carried out to obtain the ULFO mode. Finally, the reward can be calculated according to the equation (25). Based on the equation (30)-(34), the parameters of networks can be updated. With the increasing of the episodes, the training process of network converges.

Fig. 7 shows that the average reward (The average reward of each episode is the average value over these 100 evaluation episodes). The mean and the standard deviation of the average reward are indicated by the solid lines and the shaded areas, respectively. It can be seen from Fig. 7 that, during the initial stage of training process, the reward is lower. However, with the increase of training episodes, more experiences are accumulated and the reward increases rapidly. After about 4500 episodes, it converges to 0, which means that the agent has successfully learned the control policy to make the ULFO mode with satisfied damping during various operating conditions.

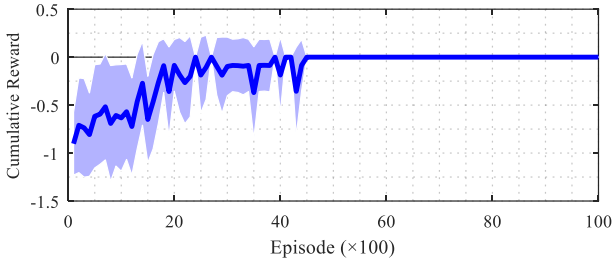


Fig. 7. Cumulative reward change with episode during the training process.

C. On-line application of the proposed agent

After off-line training, the well-trained agent can be applied for PR-PSS parameter settings self-tuning. The following online application strategy is proposed for the agent and the details are shown in Fig. 8.

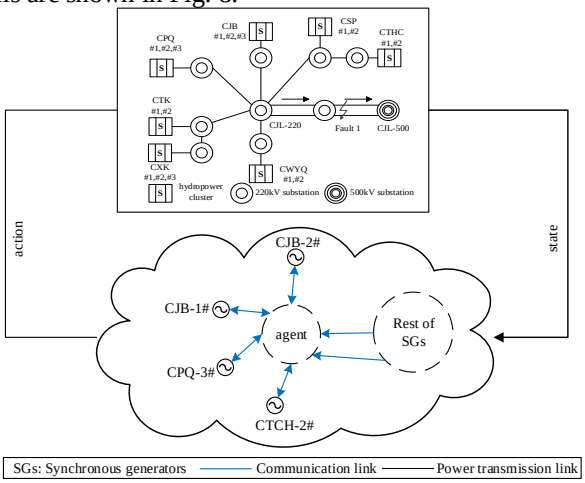


Fig. 8. Actual implementation architecture of proposed agent.

Step 1: Based on the dynamic trajectory of the system (e.g. the power angle the generator), the eigenvalues of the system can be identified by Prony algorithm.

Step 2: Analysis of the identified eigenvalues to judge whether the system has the risk of ULFO. If not, the above steps are repeated; if yes, the strategy proceeds to the next step.

Step 3: Upload the hydraulic turbine operating conditions information. The agent provides the optimal parameter settings for the hydraulic turbine equipped with PR-PSS based on the feedback information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

D. Performance Evaluation of PR-PSS

In order to illustrate the effectiveness of the proposed method, the range of T_w is discretized to form several typical operating conditions, which is shown in Table III.

TABLE III
Values of T_w in different cases

Cases	Case 1	Case 2	Case 3	Case 4	Case 5
T_w	0.5	1	1.5	2	2.5

Taking Case 1 as an example, a two-phase short circuit fault, which starts at 2 s and lasts 100 ms (Fault 1 in Fig. 6) is applied to the external line of the CJL-220KV substation in the test system. The frequency deviations of partial generators are shown in Fig. 9. It can be seen that the original system (the system is not equipped with PSS and PR-PSS) is unstable and the frequencies of the units are constantly oscillating for a long period. However, by applying the PR-PSS in the corresponding units, the frequency fluctuation of each unit is obviously reduced and quickly restored to the steady state. Therefore, it can be concluded that the proposed PR-PSS can suppress ULFO.

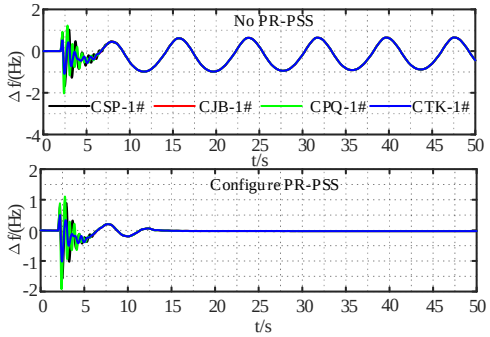


Fig. 9. Frequency deviation subject to fault without and with PR-PSS.

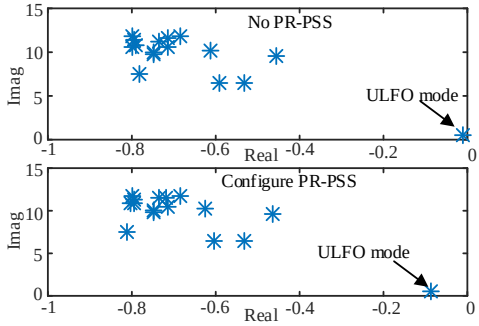


Fig. 10. Eigenvalue distribution of systems without and with PR-PSS.

To further illustrate the changes in the system stability after applying the PR-PSS, a distribution of eigenvalues of the test system (Case 1) is shown in Fig. 10. It can be seen that only the ULFO mode moves left, and the other eigenvalues have hardly changed. That is to say, the PR-PSS only enhances the damping of the ULFO mode and has less effect on the other modes, which is consistent with the conclusions of Section IV-A.

To compare PR-PSS with other ULFO suppression strategies, two common strategies are also used as the cases:

Optimizing PID parameters: suppression of ULFO by optimizing governor PID parameters of hydro-turbine units, where the more details can be found in [7].

Applying GPSS: GPSS are configured in CTHC-2#, CPQ-3#, CJB-1#, CJB-2# and CJB-3#, and the parameters of GPSS are tuned based on [12].

So far, three methods of ULFO suppression have been proposed, and applied to test system (Case 1). Frequency deviations of CTK-1# are shown in Fig. 11 and the change of ULFO mode is listed in Table IV.

TABLE IV
Comparison of ULFO modes by using different strategies

Strategies	Real part (α)	Imag part (σ)	Damping (ξ)
Original settings	-0.012	0.570	2.10%
Applying GPSS	-0.034	0.535	6.34%
Optimizing PID parameters	-0.053	0.603	8.78%
Applying PR-PSS	-0.054	0.562	9.56%

It can be seen from Fig. 11 and Table IV that all three strategies can effectively suppress ULFO. Moreover, by comparing the damping change of UFLO modes under three methods, it can be concluded that PR-PSS has a better enhancement effect on the ULFO mode damping than the optimizing governor parameters and applying GPSS.

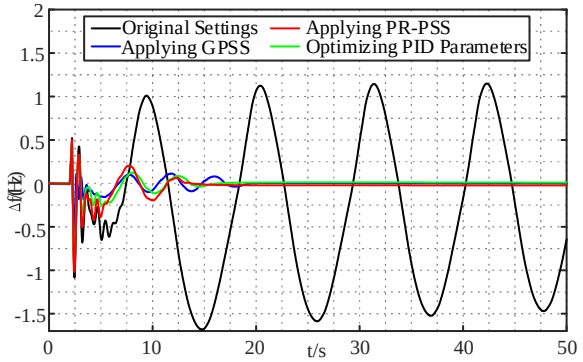


Fig. 11. Frequency deviation of CTK-1# with different damping strategies.

E. Robustness Test of the Proposed A3C-Based Agent

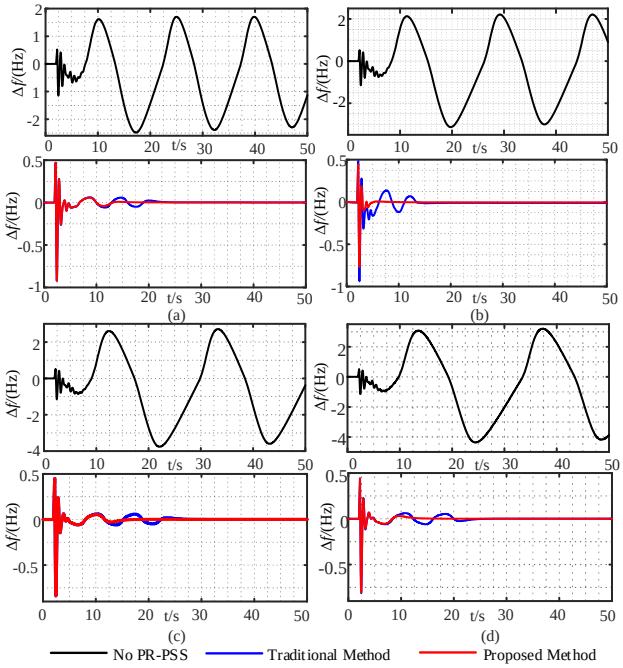


Fig. 12. Frequency deviation of CTK-1# during Fault 1: (a) Case 2; (b) Case 3; (c) Case 4; (d) Case 5. (see Table II)

In this paper, an A3C-based agent is proposed for adaptive adjustment of PR-PSS parameters settings. To evaluate the adaptability and robustness of the proposed agent, the four cases in Table II are taken as comparison scenarios. In each case, the trained agent provides the corresponding parameter settings for PR-PSS. The PR-PSS tuning based on the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

traditional method is also performed for comparison. (more details can be found in [15]). Time domain simulations are carried out in these four cases and the dynamic response of the CTK-1# is shown in Fig. 12.

According to Fig. 12, during four cases, the system with the proposed A3C-based agent reaches steady state in the shortest time. Compared with the traditional method, it shows better performance in suppressing ULFO. Therefore, it can be concluded that the robustness of the proposed method is much better than the traditional methods.

To further verify the robustness of the proposed method, the time constant T_w of all turbines are selected randomly in range [0.5, 2.5]. Each set of parameters forms a separate case. For all cases (2187 cases in total), the ULFO modes and its probability density function (PDF) are calculated and plotted in the complex plane, which are shown in Fig. 13.

It can be seen from the Fig. 13 (a) that the ULFO mode are located in the right plane of real axis in some cases, which means that the system has a risk of ULFO. Fig. 13 (b) and (c) shows the results of ULFO mode with the traditional method and the proposed method, respectively. In all cases, the ULFO modes are located in the left plane of the real axis. Moreover, compared with the traditional method, the proposed method makes the ULFO mode to move more to the left, and the system has a larger stability margin.

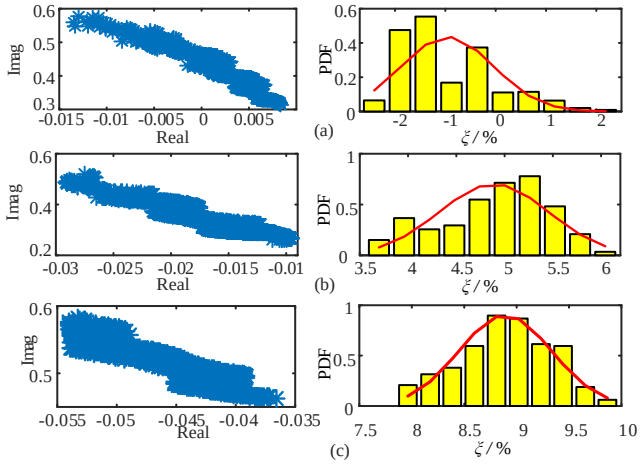


Fig. 13. ULFO modes in different conditions: (a) No PR-PSS; (b) PR-PSS tuned with traditional method; (c) PR-PSS tuned with proposed method.

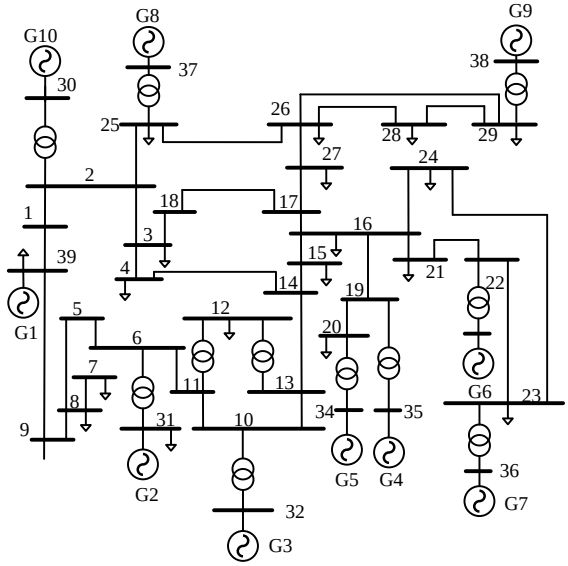


Fig. 14. Transmission network for 10-machine 39-bus system.

VI. FURTHER DISCUSSION IN IEEE BENCHMARK MODEL

A. IEEE 10-machine 39-bus system

According to [26], the IEEE 10-machine 39-bus system can be used as an benchmark model for the analysis and control of small-signal stability. Therefore, this paper also adopts this model as a test system to further verify the effectiveness of the proposed method. The structure of this benchmark model is shown in Fig. 14.

TABLE V
Values of T_w in different cases

Cases	Case 1	Case 2	Case 3
T_w	1	2	2.5

B. Comparison of different reward function

For the test system, Fault 1 (Three-phase short circuit fault occurring in the bus 5, starting at 2.0 s and lasting for 0.2 s.) is used to excite the dynamic characteristics of the system, and the Prony method [27] is used to identify the eigenvalues of the system. The results show that there is an ULFO mode 0.002+j0.56 in the test system, and it would make the test system to have high risk of ULFO. To solve this problem, PR-PSS are configured on the G1, G5 and G8, and the A3C algorithm is used to train the agent for the PR-PSS parameters self-tuning.

For the DRL, the selection of reward is very important. Since the reward is used to evaluate the merits of the action. If the design of reward function is unreasonable, it also affects the convergence speed of the algorithm.

For the PR-PSS parameters self-tuning problem, four methods can be used to form reward function: mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE). These indicators are computed as follows:

$$
\begin{aligned}
MAE &= \frac{1}{n}\sum_{i=1}^{n}\sum_{\xi_i \le \xi_{set}} \left|\xi_i - \xi_{set}\right| \\
MSE &= \frac{1}{n}\sum_{i=1}^{n}\sum_{\xi_i \le \xi_{set}} \left(\xi_i - \xi_{set}\right)^2 \\
RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sum_{\xi_i \le \xi_{set}} \left(\xi_i - \xi_{set}\right)^2} \\
MAPE &= \frac{1}{n}\sum_{i=1}^{n}\sum_{\xi_i \le \xi_{set}} \frac{\left|\xi_i - \xi_{set}\right|}{\xi_{set}}
\end{aligned}
\tag{35}
$$

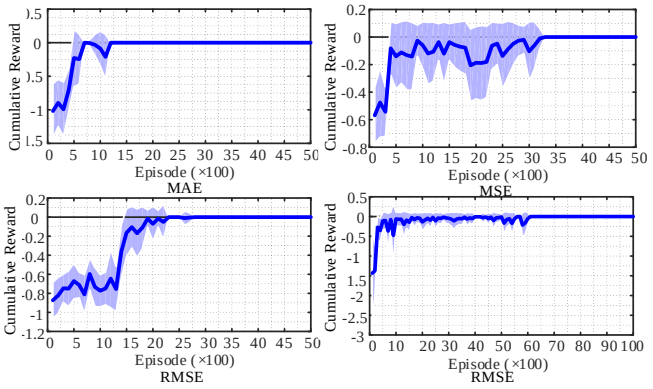


Fig. 15. Cumulative reward change with episode during the training process.

Fig. 15 shows that the average reward (The average reward of each episode is the average value over these 100 evaluation episodes). The mean and the standard deviation of the average reward are indicated by the solid lines and the shaded areas,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2020.2997790, IEEE Transactions on Smart Grid

respectively. It can be seen from Fig. 15 that using the reward function constructed by MAE to train the agent can get the best training effect, and it makes the algorithm to converge fastest.

C. Stability guarantees of the proposed agent

To ensure the stability of the proposed agent, this paper propose an strategy for the agent: Firstly, a set of conservative PR-PSS parameter setting can be pre-calculated by the robust method, and it can still guarantee the system stability in extreme scenarios. Then, during each state, the parameter setting provided by the well-trained agent can be compared with the pre-calculated conservative parameter setting in the estimation system. Finally, the better one (which obtains the bigger reward) can be sent to the actual system. The detailed process is shown in Fig. 16.

Estimation system: It can be seen as the dynamic equivalent model of the actual system, and it would update the parameters of the system based on the deviation of state estimation to approximate the actual system [17].

Comparison device: Based on the estimation system, the action provided by the agent is compared with the pre-calculated conservative action, and the better one (which obtains the bigger reward) is sent to the actual system.

Step1: In each state, based on the feedback information, the well-trained agent will provide an action (PR-PSS parameter settings) to the comparison device.

Step2: The comparison device will send the action provided by the agent and pre-calculated action to the estimation system, respectively. Then, the estimation system would feedback the corresponding rewards.

Step3: The comparison device will judge the two actions based on the rewards, and send the better one to the actual system.

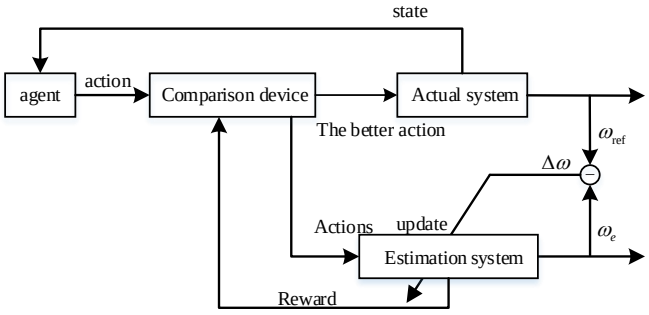


Fig. 16. A3C-based agent training diagram to be applied online.

In this way, we can indirectly limit the behavior of the controller, thus avoiding the controller could have unusual, unexpected behavior in particular scenarios.

D. Comparison of the proposed agent with other methods

In order to compare PR-PSS with other ULFO suppression strategies, two other very well classical PSS tuned devices are also used as the test cases:

Configure GPSS: GPSSs are equipped in G1, G5 and G8, and the parameters of GPSSs are tuned based on pole-placement method. More details can be found in [12];

Configure PSS4B: PSS4Bs are equipped in G1, G5 and G8, and the parameters of PSS4Bs are tuned based on [14];

It can be seen from Fig. 17 that PR-PSS shows better performance in suppressing ULFO in Case 1 when compared with GPSS and PSS4B. Therefore, it can be concluded that the effeteness of the proposed method is much better than the other classical PSS tuned devices.

TABLE VI
Comparison of ULFO modes during Case1

Case	Strategies	Real part (α)	Imag part (σ)	Damping (ξ)
Case 1	Original settings	0.002	0.57	-0.35%
	PSS4B	-0.022	0.56	3.89%
	GPSS	-0.038	0.57	6.61%
	PR-PSS	-0.063	0.58	10.80%

Moreover, to further compare the proposed method with PSS4B and GPSS, a distribution of ULFO modes of the test system is shown in Table VI. It can concluded that the proposed method can make the ULFO mode to move more to the left, and the system has a larger stability margin.

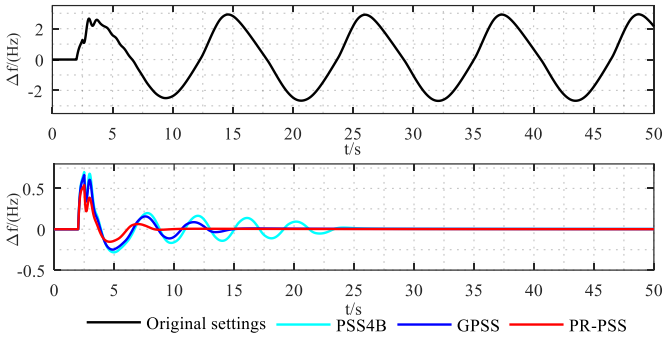


Fig.17. Frequency deviation of G1 during Case 1.

It can be observed from Figs. 18-19 that for Case 2 and Case 3, the system with the proposed A3C-based tuning strategy reaches steady state in the shortest time compared with the PSS4B and GPSS. It means that the proposed method has a better robustness than the other two methods in the cases with uncertainties.

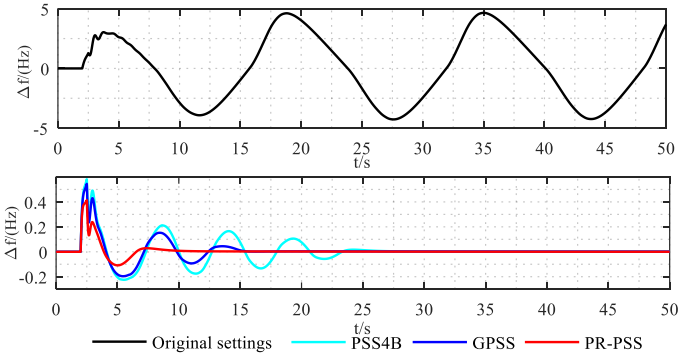


Fig.18. Frequency deviation of G1 during Case 2.

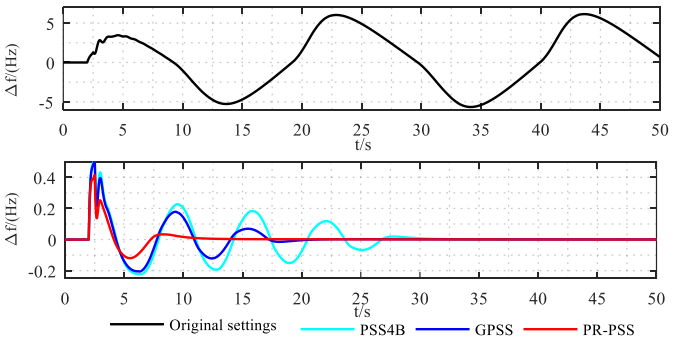


Fig.19. Frequency deviation of G1 during Case 3.

VII. CONCLUTION

In this paper, a PR-PSS is proposed and designed to prevent ULFO in the power system, and an A3C-based agent is proposed for adaptive adjustment of PR-PSS parameters settings. Simulation results show that the PR-PSS can effectively suppress ULFO and has less negative effect on other oscillation modes. Compared with other damping strategies, applying PR-PSS shows better performance in suppressing ULFO. Both time domain simulations and eigenvalue analysis results demonstrate that the PR-PSS parameter settings tuned by the proposed agent work well and show much better performance compared to the traditional method.

REFERENCES

[1] R. Xie, I. Kamwa, D. Rimorov, et al. "Fundamental study of common mode small-signal frequency oscillations in power systems," *Int. J. Electr. Power Energy Syst.*, vol. 106, pp. 201-209, Mar. 2019.

[2] H. Villegas Pico, J. D. McCalley, A. Angel, et al. "Analysis of Very Low Frequency Oscillations in Hydro-Dominant Power Systems Using Multi-Unit Modeling," *IEEE Trans. Power Syst*., vol. 27, no. 4, pp. 1906-1915, Nov. 2012.

[3] Z. Liu, W. Yao, J. Wen, et al. "Effect analysis of generator governor system and its frequency mode on inter-area oscillations in power systems," *Int. J. Electr. Power Energy Syst.*, vol. 96, pp. 1-10, Mar. 2018.

[4] G. Chen, F. Tang, H. Shi, et al. "Optimization Strategy of Hydro-governors for Eliminating Ultra-low-Frequency Oscillations in Hydro-dominant Power Systems," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 3, pp. 1086-1094, Sept. 2018.

[5] Y. Chen, Y. Liu, Z. Tang et al. "Analysis of ultra-low frequency oscillation in yunnan asynchronous sending system," 2017 *IEEE Power & Energy Society General Meeting*, Chicago, IL, 2017, pp. 1-5.

[6] W. Mo, Y. Chen, H. Chen, et al. "Analysis and Measures of Ultra-low-Frequency Oscillations in a Large-Scale Hydropower Transmission System," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 3, pp. 1077-1085, Sept. 2018.

[7] L. Chen, X. Liu, Y. Min, et al. "Optimization of Governor Parameters to Prevent Frequency Oscillations in Power Systems," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4466-4474, July 2018.

[8] H. N. Villegas. "Electromechanical Oscillations in Hydro-Dominant Power Systems: An Application to the Colombian Power System," (2011). Graduate Theses and Dissertations. Paper 10116.

[9] C. Jiang, J. Zhou, P. Shi, et al. "Ultra-low frequency oscillation analysis and robust fixed order control design," *Int. J. Electr. Power Energy Syst.*, vol. 104, pp. 269-278, Jan. 2019.

[10] L. Chen, Y. Min, X. Lu, et al. "Online emergency control to suppress frequency oscillations based on damping evaluation using dissipation energy flow, " *Int. J. Electr. Power Energy Syst.*, vol. 103, pp. 414-420, Dec. 2018.

[11] R. Grondin, I. Kamwa, L. Soulieres, et al. "An approach to PSS design for transient stability improvement through supplementary damping of the common low-frequenc," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 954-963, Aug. 1993.

[12] S. Liu, D. Wang, N. Ma, et al. " Study on Characteristics and Suppressing Countermeasures of Ultra-low Frequency Oscillation Caused by Hydropower Units, " *Proceedings of the CSEE*, vol. 18, pp. 5354-5362, 2019 (in Chinese).

[13] F. R. Schleif, G. E. Martin and R. R. Angell. "Damping of System Oscillations with a Hydrogenerating Unit," *IEEE Trans. Power App. Syst*, vol. 86, no. 4, pp. 438-442, April 1967.

[14] I. Kamwa, R. Grondin and G. Trudel, "IEEE PSS2B versus PSS4B: the limits of performance of modern power system stabilizers," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 903-915, May 2005.

[15] J. Zhang, C. Y. Chung, C. Lu, et al. "A Novel Adaptive Wide Area PSS Based on Output-Only Modal Analysis," *IEEE Trans. Power Syst*., vol. 30, NO. 5, PP. 2633-2642, SEPT. 2015.

[16] M. A. Abido. "Optimal design of power-system stabilizers using particle swarm optimization," *IEEE Trans. Energy Convers.*, vol. 17, no. 3, pp. 406-413, Sept. 2002.

[17] P. Hoang and K. Tomsovic. "Design and analysis of an adaptive fuzzy power system stabilizer, " *IEEE Trans. Energy Convers.*, vol. 11, no. 2, pp. 455-461, June 1996.

[18] T. C. Yang. "Applying H ∞ optimisation method to power system stabiliser design Part 2: Multi-machine power systems." *Int. J. Electr. Power Energy Syst*., vol. 19, no. 1, pp. 37-43, Jan. 1997.

[19] R. Jalayer and B. Ooi. "Co-Ordinated PSS Tuning of Large Power Systems by Combining Transfer Function-Eigenfunction Analysis (TFEA), Optimization, and Eigenvalue Sensitivity," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2672-2680, Nov. 2014.

[20] R. Hadidi and B. Jeyasurya. "Reinforcement Learning Based Real-Time Wide-Area Stabilizing Control Agents to Enhance Power System Stability," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 489-497, Mar. 2013.

[21] V. Mnih, A. Badia, M. Mirza, et al. "Asynchronous Methods for Deep Reinforcement Learning, " *2016 International Conference on Machine Learning*, New York, 2016, pp. 1-10.

[22] H. Hua, Y. Qin, C. Hao, et al. "Optimal energy management strategies for energy Internet via deep reinforcement learning approach," *Applied Energy*, vol. 239, pp. 598-609, Apr. 2019.

[23] H. Wang. "A unified model for the analysis of FACTS devices in damping power system oscillations. III. Unified power flow controller," *IEEE Trans. Power Del*., vol. 15, no. 3, pp. 978-983, July 2000.

[24] W. Chen and J. Lin. "One-Dimensional Optimization for Proportional–Resonant Controller Design Against the Change in Source Impedance and Solar Irradiation in PV Systems," *IEEE Trans. Ind. Electron.*, vol. 61, no. 4, pp. 1845-1854, April 2014.

[25] Z. Wan, H. Li, H. He, et al "Model-Free Real-Time EV Charging Scheduling Based on Deep Reinforcement Learning," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5246-5257, Sept. 2019.

[26] S. Canizares T. Fernandes, E. Geraldi, et al. "Benchmark Models for the Analysis and Control of Small-Signal Oscillatory Dynamics in Power Systems," *IEEE Trans. Power Syst*., vol. 32, no. 1, pp. 715-722, Jan. 2017.

[27] G. Wang, Z. Xu, X. Guo, et al. "Mechanism analysis and suppression method of ultra-low-frequency oscillations caused by hydropower units," *Int. J. Electr. Power Energy Syst*., vol. 103, pp. 102-114, Dec. 2018.

APPENDIX

For the hydropower system in SCPG, the parameters of GPSS, PSS4B and PR-PSS are shown in the following tables.

Table A. Parameters of the PID

	Units	K_p	K_I	K_D
PID	CJB-1#	1.21	0.34	2.47
	CJB-2#	2.90	0.57	2.03
	CPQ-3#	3.72	0.69	4.10
	CTCH-3#	3.58	0.33	3.76

Table B. Parameters of the GPSS

	Units	K_{STAB}	T_1	K_3
GPSS	CJB-1#	48.56	0.46	0.59
	CJB-2#	22.35	0.32	0.70
	CPQ-3#	30.48	0.05	0.55
	CTCH-3#	29.12	0.18	0.61

Table C. Parameters of the PR-PSS (only for Case1)

	Units	Parameters				
PR-PSS	CJB-1#	K_p	K_R	w_c	w_0	a_0
		0.21	0.79	0.09	0.57	5.59
		a_1	a_2	b_0	b_1	b_2
		7.89	6.41	3.81	4.25	5.04
	CJB-2#	K_p	K_R	w_c	w_0	a_0
		0.21	0.79	0.09	0.57	5.36
		a_1	a_2	b_0	b_1	b_2
		4.17	7.12	6.82	6.27	3.75
	CPQ-3#	K_p	K_R	w_c	w_0	a_0

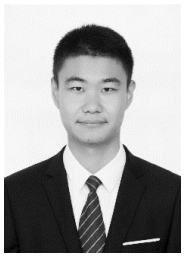
	0.21	0.79	0.09	0.57	4.23
	a_1	a_2	b_0	b_1	b_2
	6.14	3.77	8.49	5.44	2.17
	K_p	K_R	w_c	w_0	a_0
CTCH-3#	0.21	0.79	0.09	0.57	3.28
	a_1	a_2	b_0	b_1	b_2
	4.79	3.90	6.11	5.38	8.55

Guozhou Zhang received the B.S. from Chongqing University of Technology, Chongqing, China, in 2016, the M. S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2019. He is currently working toward the Ph.D. degree in control science and engineering at the UESTC. His research interest includes power system anlysis and control.

Weihao Hu (S'06–M'13–SM'15) received the B.Eng. and M.Sc. degrees from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively, both in electrical engineering, and Ph. D. degree from Aalborg University, Denmark, in 2012.

He is currently a Full Professor and the Director of Institute of Smart Power and Energy Systems (ISPES) at the University of Electronics Science and Technology of China (UESTC). He was an Associate Professor at the Department of Energy Technology, Aalborg University, Denmark and the Vice Program Leader of Wind Power System Research Program at the same department. His research interests include artificial intelligence in modern power systems and renewable power generation. He has led/participated in more than 15 national and international research projects and he has more than 170 publications in his technical field.

He is an Associate Editor for IET Renewable Power Generation, a Guest Editor-in-Chief for Journal of Modern Power Systems and Clean Energy Special Issue on Applications of Artificial Intelligence in Modern Power Systems, a Guest Editor-in-Chief for Transactions of China Electrical Technology Special Issue on Planning and operation of multiple renewable energy complementary power generation systems, and a Guest Editor for the IEEE TRANSACTIONS ON POWER SYSTEM Special Section on Enabling very high penetration renewable energy integration into future power systems. He was serving as the Technical Program Chair (TPC) for IEEE Innovative Smart Grid Technologies (ISGT) Asia 2019 and is serving as the Conference Chair for the Asia Energy and Electrical Engineering Symposium (AEEES 2020). He is currently serving as Chair for IEEE Chengdu Section PELS Chapter and he is an IEEE Senior Member.



Di Cao received the B.S. from North China University of Water Resources and Electric Power, Zhengzhou, China, in 2014. He is currently working toward the Ph.D. degree in control science and engineering at the University of Electronic Science and Technology of China. His research interest includes optimization of distribution network.

Qi Huang (S'99, M'03, SM'09) was born in Guizhou province in the People's Republic of China. He received his BS degree in Electrical Engineering from Fuzhou University in 1996, MS degree from Tsinghua University in 1999, and Ph.D. degree from Arizona State University in 2003. He is currently a professor at UESTC, the Executive Dean of School of Energy Science and Engineering, UESTC, and the director of Sichuan State Provincial Lab of Power System Wide-area Measurement and Control. He is a member of IEEE since 1999. His current research and academic interests include power system instrumentation, power system monitoring and control, and power system high performance computing.

Jianbo Yi (S' 08-M' 12) was born in Gansu province, China. He received his BS, MS and Ph. D degree in Electrical Engineering from UESTC in 2004, 2007, and 2013 respectively. He is currently an associate professor at School of Mechanical and Electrical Engineering, UESTC. His research interests are power system analysis and control.

Zhe Chen (M'95-SM'98-F'19) received the B.Eng. and M.Sc. degrees from Northeast China Institute of Electric Power Engineering, Jilin, China, and the Ph.D. degree from University of Durham, Durham, U.K.

He is a Full Professor with the Department of Energy Technology, Aalborg University, Denmark. He is the Leader of Wind Power System Research Program in the Department of Energy Technology, Aalborg University and the Danish Principle Investigator for Wind Energy of Sino-Danish Centre for Education and Research. His research areas include power systems, power electronics and electric machines; and his main current research interests are wind energy and modern power systems. He has led many research projects and has more than 400 publications in his technical field.

Dr. Chen is an Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, an Associate Editor of the IEEE TRANSACTIONS ON POWER ELECTRONICS, a Fellow of the Institution of Engineering and Technology, London, U.K., a Chartered Engineer in the U.K. A Fellow of the IEEE.

Frede Blaabjerg (S'86–M'88–SM'97–F'03) was with ABB-Scandia, Randers, Denmark, from 1987 to 1988. From 1988 to 1992, he got the PhD degree in Electrical Engineering at Aalborg University in 1995. He became an Assistant Professor in 1992, an Associate Professor in 1996, and a Full Professor of power electronics and drives in 1998. From 2017 he became a Villum Investigator. He is honoris causa at University Politehnica Timisoara (UPT), Romania and Tallinn Technical University (TTU) in Estonia.

His current research interests include power electronics and its applications such as in wind turbines, PV systems, reliability, harmonics and adjustable speed drives. He has published more than 600 journal papers in the fields of power electronics and its applications. He is the co-author of four monographs and editor of ten books in power electronics and its applications.

He has received 32 IEEE Prize Paper Awards, the IEEE PELS Distinguished Service Award in 2009, the EPE-PEMC Council Award in 2010, the IEEE William E. Newell Power Electronics Award 2014, the Villum Kann Rasmussen Research Award 2014, the Global Energy Prize in 2019 and the 2020 IEEE Edison Medal. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON POWER ELECTRONICS from 2006 to 2012. He has been Distinguished Lecturer for the IEEE Power Electronics Society from 2005 to 2007 and for the IEEE Industry Applications Society from 2010 to 2011 as well as 2017 to 2018. In 2019-2020 he serves as President of IEEE Power Electronics Society. He is Vice-President of the Danish Academy of Technical Sciences too. He is nominated in 2014-2019 by Thomson Reuters to be between the most 250 cited researchers in Engineering in the world.